

AD-A013 068

ESTIMATION OF PARAMETERS IN SYSTEMS OF  
RELATIVELY SLOW TIME VARIATION

Philip H. Fiske, et al

Michigan University

Prepared for:

Air Force Office of Scientific Research

October 1974

DISTRIBUTED BY:

**NTIS**

National Technical Information Service  
U. S. DEPARTMENT OF COMMERCE

224101

ESTIMATION OF PARAMETERS IN SYSTEMS OF RELATIVELY SLOW TIME VARIATION

Philip H. Fiske  
William L. Root

The University of Michigan

Approved for public release;  
distribution unlimited.

Presented at

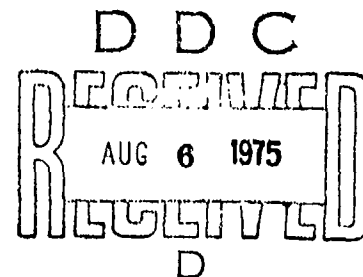
Ninth Annual Allerton Conference on Circuit and System Theory

October 1974

Reproduced by  
NATIONAL TECHNICAL  
INFORMATION SERVICE  
US Department of Commerce  
Springfield, VA. 22151

Research sponsored by the

United States Air Force  
Air Force Office of Scientific Research  
Grant No. 72-2328-B



AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)  
NOTICE OF TRANSMITTAL TO DDC  
This technical report has been reviewed and is  
approved for public release IAW AFR 190-12 (7b).  
Distribution is unlimited.

D. W. TAYLOR  
Technical Information Officer

ADA013068

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR - TR - 75 - 1033	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ESTIMATION OF PARAMETERS IN SYSTEMS OF RELATIVELY SLOW TIME VARIATION		5. TYPE OF REPORT & PERIOD COVERED Interim
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Philip H. Fiske William L. Root		8. CONTRACT OR GRANT NUMBER(s) AFOSR 72-2328
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Michigan Aerospace Engineering Department Ann Arbor, MI 48105		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 9749-06
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research (NM) 1400 Wilson Blvd. Arlington, Virginia 22209		12. REPORT DATE October 1974
		13. NUMBER OF PAGES 8 12
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES Twelfth Allerton Conference on Circuit and System Theory October 1974, University of Illinois		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Parameter estimation. Nonlinear, time-varying system. Identification.		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) The estimation of parameters for a rather general class of nonlinear, time-varying, causal, bounded memory systems is discussed. A model for such systems is established using the concept that the system variation is unknown, but bounded between observation times. A recursive estimation algorithm is developed for the model, and both upper and lower bounds are found for the guaranteed error of estimation.		

# ESTIMATION OF PARAMETERS IN SYSTEMS OF RELATIVELY SLOW TIME VARIATION\*

Philip H. Fiske  
William L. Root  
The University of Michigan, Ann Arbor

## ABSTRACT

The estimation of parameters for a rather general class of nonlinear, time-varying, causal, bounded memory systems is discussed. A model for such systems is established using the concept that the system variation is unknown, but bounded between observation times. A recursive estimation algorithm is developed for the model, and both upper and lower bounds are found for the guaranteed error of estimation.

## DEVELOPMENT OF THE SYSTEM MODEL

In this paper the term system means simply a mapping from an appropriate space of inputs to an appropriate space of outputs. The central purpose of the paper is to obtain estimates on how rapidly causal systems, of a certain rather general class that includes nonlinear as well as linear systems, can vary with time and still allow adequate experimental identification.

A particular kind of system model is especially appropriate for this study and is used here. It has the features that it is linear in the parameters, even though the system itself may be nonlinear, and that it is widely applicable if the only noise to be considered is output observation noise. We briefly and heuristically describe such a model in a special case in what follows, and this is sufficient for this paper. However there is a general theory of such models, and most of the ad hoc assumptions made below can either be shown to be justifiable or not needed (see [1]).

Let  $\mathcal{X}$  be an arbitrary input space,  $\mathcal{Y}$  a linear output space, and  $\mathcal{H}$  a set of mappings from  $\mathcal{X}$  into  $\mathcal{Y}$  (i.e.,  $\mathcal{H}$  is a class of "systems"), so that one may write

$$y = h(x), \quad x \in \mathcal{X}, \quad y \in \mathcal{Y}, \quad h \in \mathcal{H}. \quad (1)$$

Each  $x \in \mathcal{X}$  determines a function from  $\mathcal{H}$  into  $\mathcal{Y}$ . If this function is denoted by  $X$ , then

$$y = X(h), \quad h \in \mathcal{H}. \quad (2)$$

Let addition and scalar multiplication be defined in  $\mathcal{H}$  in the way they usually are for functions and then extend  $\mathcal{H}$ , if necessary, so that it is closed under linear combinations. Denote this linear extension by  $\mathcal{F}$ . It is now trivial to verify that the mapping  $X$  as extended to  $\mathcal{F}$  will be linear, even

---

\*Research sponsored by the United States Air Force, Air Force Office of Scientific Research Grant No. 72-2328-B.

though the mapping  $h$  may be nonlinear (see e.g., [1]). If additive output noise is present, (1) and (2) are replaced by

$$y = h(x) + v \quad (3)$$

and

$$y = Xh + v, \quad h \in \mathcal{F}, \quad (4)$$

where  $v$  is a random variable taking values in  $\mathcal{Y}$  and  $X$  is linear. Thus the problem of identifying the system has been put in the form of the classical problem of estimating  $h$  in the linear model (4). Of course, equation (4) is abstract, so we now need to specialize it so as to have a meaningful estimation problem. Here we simply assume what we need in order to do this easily.

Suppose that a system with vector-valued inputs and outputs can be adequately represented in terms of sampled inputs and outputs. Furthermore, assume that it has bounded memory (this is the only critical assumption). Then an input is a vector sequence  $\{x_k\}$ , the corresponding output is a vector sequence  $\{y_k\}$ , and if the system has memory  $m$ ,

$$\begin{aligned} y_k &= h_k(x_k, x_{k-1}, \dots, x_{k-m}) + v_k \\ \text{or} \quad y_k &= h_k(\underline{x}_k) + v_k, \quad k = 1, 2, \dots, \\ \text{where} \quad \underline{x}_k &\triangleq (x_k, x_{k-1}, \dots, x_{k-m}). \end{aligned} \quad (5)$$

If the system is time-invariant each  $h_k$  is the same; otherwise, of course, the  $h_k$ 's are different. To each  $\underline{x}_k$  there corresponds a linear mapping  $X_k$  defined, as before, by  $X_k h_k \triangleq h_k(\underline{x}_k)$ , so that

$$y_k = X_k h_k + v_k, \quad k = 1, 2, \dots \quad (6)$$

Let there be imposed the further condition that all the mappings  $h_k \in \mathcal{F}$  are fully characterized by a finite set of parameters, say  $p$  in number. The  $\{h_k\}$  can then be represented by  $p$ -vectors, and if the output is an  $r$ -vector, the  $\{X_k\}$  can be represented by  $r \times p$  matrices.

If the systems in  $\mathcal{F}$  are of uniformly bounded time variation for each finite interval, one can write

$$h_{k+1} = h_k + w_k, \quad k = 1, 2, \dots, \quad (7)$$

where each component of  $w_k$  is bounded. The system model then becomes

$$\begin{cases} h_{k+1} = h_k + w_k \\ y_k = X_k h_k + v_k, \quad k = 1, 2, \dots, \end{cases} \quad (8)$$

where the  $\{w_k\}$  are bounded, unknown elements of  $R^p$  and the  $\{v_k\}$  are random variables representing output noise. The equations (8) are, of course, in the form of the system equations for a degenerate linear dynamical system, but the interpretation given to them is entirely different.

We wish to consider successive blocks of measurements where each block

consists of tests with the same suitably chosen  $N$  inputs  $(x_1, \dots, x_N)$ . We make the important simplifying assumption that the system parameters change only between blocks of measurements. Results can be obtained without this assumption (see [3]), but there is not space to go into the additional argument necessary here. This assumption gives an acceptable approximation anyway in many cases where there is only slight variation within the blocks.

With this assumption we have

$$\begin{cases} h^{k+1} = h^k + w^k \\ y^k = Xh^k + v^k, \quad k = 1, 2, \dots, \end{cases} \quad (9)$$

where  $h^k \triangleq h_{kN}$  and

$$y^k = \begin{bmatrix} y_{(k-1)N+1} \\ y_{(k-1)N+2} \\ \vdots \\ y_{kN} \end{bmatrix}, \quad X \triangleq \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}, \quad v^k = \begin{bmatrix} v_{(k-1)N+1} \\ v_{(k-1)N+2} \\ \vdots \\ v_{kN} \end{bmatrix}.$$

The condition to be imposed on the inputs  $(x_1, \dots, x_N)$  is that the operator  $X$  has zero null space, so that the vector  $h^k$  is estimable. The usual assumptions are made about the noise,

$$E v^k = 0, \quad E v^k (v^j)^T = R \delta_{jk}, \quad (10a)$$

where  $R$  is strictly definite. The boundedness condition on the  $w^k$  is taken to be in the form,

$$|(w^k, \psi_i)| \leq \eta_k, \quad i = 1, \dots, p, \quad (10b)$$

for all  $k = 1, 2, \dots$ , where  $\{\psi_i\}$  is a complete orthonormal basis of eigenvectors of  $(X^T R^{-1} X)^{-1}$ , i.e.,

$$(X^T R^{-1} X)^{-1} \psi_i = \sigma_i^2 \psi_i, \quad i = 1, \dots, p.$$

The eigenvalues  $\{\sigma_i^2\}$  are real and nonnegative.

### ESTIMATION OF PARAMETERS

We now derive a recursive estimation algorithm for the parameter vector  $h^k$  and obtain a uniform upper bound on the mean-squared error. Since the actual mean-squared error depends on the  $\{w^k\}$ , which are unknown, an exact expression for mean-squared error cannot be given, of course.

We begin by forming the standard linear, unbiased, minimum variance (LUMV) estimate of  $h^1$

$$\hat{h}^1 = C y^1, \quad (11)$$

where  $C \triangleq (X^T R^{-1} X)^{-1} X^T R^{-1}$ . The error in this estimate satisfies

$$E\|\hat{h}^1 - h^1\|^2 = \sum_{i=1}^p \sigma_i^2 = \sum_{i=1}^p b_{1i}^2 \quad (12)$$

where  $b_{1i}^2 \triangleq \sigma_i^2$ ,  $i = 1, \dots, p$ . Now consider the second observation

$$y^2 = Xh^2 + v^2 = Xh^1 + Xw^1 + v^2,$$

which can be rewritten as

$$y^2 - Xh^1 = X[(h^1 - \hat{h}^1) + w^1] + v^2. \quad (13)$$

Define  $z^2 \triangleq y^2 - Xh^1$  and  $\xi^1 \triangleq (h^1 - \hat{h}^1) + w^1 = h^2 - \hat{h}^1$  so that (13) becomes

$$z^2 = X\xi^1 + v^2. \quad (14)$$

Let  $\hat{\xi}^1$  be a linear estimate of  $\xi^1$  that is a function of  $Cz^2$ , but is otherwise arbitrary. Thus,

$$\hat{\xi}^1 = \sum_{i,j=1}^p a_{ij} (Cz^2, \psi_j) \psi_i \quad (15)$$

where the  $\{a_{ij}\}$  are real numbers. The error in this estimate is given by

$$\begin{aligned} \xi^1 - \hat{\xi}^1 &= \sum_{i=1}^p \left[ (a_{ii} - 1)(\xi^1, \psi_i) + \sum_{\substack{j=1 \\ (j \neq i)}}^p a_{ij} (\xi^1, \psi_j) \right. \\ &\quad \left. + \sum_{j=1}^p a_{ij} (Cv^2, \psi_j) \right] \psi_i. \end{aligned} \quad (16)$$

Therefore, it follows from (10a) and the properties of the  $\{\psi_i\}$  that

$$\begin{aligned} E\|\hat{\xi}^1 - \xi^1\|^2 &= \sum_{i=1}^p E \left[ (a_{ii} - 1)(\xi^1, \psi_i) + \sum_{\substack{j=1 \\ (j \neq i)}}^p a_{ij} (\xi^1, \psi_j) \right]^2 \\ &\quad + \sum_{i,j=1}^p a_{ij}^2 \sigma_j^2 \leq 2 \sum_{i=1}^p (a_{ii} - 1)^2 E[(\xi^1, \psi_i)^2] \\ &\quad + 2 \sum_{i=1}^p E \left[ \sum_{\substack{j=1 \\ (j \neq i)}}^p a_{ij} (\xi^1, \psi_j) \right]^2 + \sum_{i,j=1}^p a_{ij}^2 \sigma_j^2 \end{aligned} \quad (17)$$

To minimize the upper bound in (17), one should obviously choose  $a_{ij} = 0$  for  $i \neq j$ . Making this choice yields

$$E\|\hat{\xi}^1 - \xi^1\|^2 = \sum_{i=1}^p (a_{ii} - 1)^2 E[(\xi^1, \psi_i)^2] + \sum_{i=1}^p a_{ii}^2 \sigma_i^2 \quad (18)$$

Now

$$\begin{aligned} E[(\xi^1, \psi_i)^2] &= E[(h^1 - \hat{h}^1, \psi_i)^2] + 2E[(h^1 - \hat{h}^1, \psi_i)(w^1, \psi_i)] + E[(w^1, \psi_i)^2] \\ &\leq b_{1i}^2 + 2b_{1i}\eta_i + \eta_i^2 = (b_{1i} + \eta_i)^2, \end{aligned}$$

\*Obviously  $E[(h^1 - \hat{h}^1, \psi_i)(w^1, \psi_i)] = 0$ , but in general  $E[(h^n - \hat{h}^n, \psi_i)(w^n, \psi_i)] \neq 0$ ,  $n \geq 2$ .

so

$$E\|\xi^1 - \xi^1\|^2 \leq \sum_{i=1}^p \left[ (a_{ii} - 1)^2 (b_{li} + \eta_i)^2 + a_{ii}^2 \sigma_i^2 \right]. \quad (19)$$

The upper bound in (19) is minimized by putting

$$a_{ii} = \frac{(b_{li} + \eta_i)^2}{(b_{li} + \eta_i)^2 + \sigma_i^2}, \quad i = 1, \dots, p.$$

The linear estimate that results when these values are used is

$$\xi^1 = \sum_{i=1}^p \frac{(b_{li} + \eta_i)^2}{(b_{li} + \eta_i)^2 + \sigma_i^2} (Cz^2, \psi_i) \psi_i. \quad (20)$$

Since  $h^2 = \hat{h}^1 + \xi^1$ , the estimate of  $h^2$  is taken to be

$$\hat{h}^2 = \hat{h}^1 + \xi^1. \quad (21)$$

Then

$$\begin{aligned} E\|\hat{h}^2 - h^2\|^2 &= E\|(\hat{h}^1 + \xi^1) - (\hat{h}^1 + \xi^1)\|^2 = E\|\xi^1 - \xi^1\|^2 \\ &\leq \sum_{i=1}^p \left\{ \left[ \frac{(b_{li} + \eta_i)^2}{(b_{li} + \eta_i)^2 + \sigma_i^2} - 1 \right]^2 (b_{li} + \eta_i)^2 \right. \\ &\quad \left. + \left[ \frac{(b_{li} + \eta_i)^2}{(b_{li} + \eta_i)^2 + \sigma_i^2} \right]^2 \sigma_i^2 \right\} = \sum_{i=1}^p \frac{(b_{li} + \eta_i)^2 \sigma_i^2}{(b_{li} + \eta_i)^2 + \sigma_i^2} = \sum_{i=1}^p b_{2i}^2 \end{aligned} \quad (22)$$

where

$$b_{2i}^2 \triangleq \frac{(b_{li} + \eta_i)^2 \sigma_i^2}{(b_{li} + \eta_i)^2 + \sigma_i^2}, \quad i = 1, \dots, p.$$

If this process is continually repeated, we obtain the estimation algorithm

$$\begin{aligned} \hat{h}^k &= \hat{h}^{k-1} + \xi^k \\ \xi^k &= \sum_{i=1}^p \frac{(b_{(k-1)i} + \eta_i)^2 \sigma_i^2}{(b_{(k-1)i} + \eta_i)^2 + \sigma_i^2} (Cy^k - \hat{h}^{k-1}, \psi_i) \psi_i \\ b_{ki}^2 &= \frac{(b_{(k-1)i} + \eta_i)^2 \sigma_i^2}{(b_{(k-1)i} + \eta_i)^2 + \sigma_i^2}, \quad i = 1, \dots, p. \end{aligned} \quad (23)$$

The initial conditions are given by

$$\hat{h}^1 = Cy^1, \quad b_{li}^2 = \sigma_i^2, \quad i = 1, \dots, p \quad (24)$$

and the error in the estimate of  $h^k$  satisfies

$$E\|\hat{h}^k - h^k\|^2 \leq \sum_{i=1}^p b_{ki}^2 \quad (25)$$

Lemma 1. If  $f(x)$  is twice continuously differentiable for  $x > 0$  and satisfies

$$(i) \quad \lim_{x \rightarrow 0+} f(x) > 0$$

- (ii)  $\lim_{x \rightarrow \infty} f(x) < \infty$   
 (iii)  $\frac{df(x)}{dx} > 0, x > 0$   
 (iv)  $\frac{d^2f(x)}{dx^2} < 0, x > 0$

then (a)  $f(x) = x$  has a unique solution  $x^*$  for  $x > 0$ .

(b) the sequence  $x_n = f(x_{n-1}), x_1 > 0$ , converges to  $x^*$ .

Proof: Follows from elementary arguments. ||| The limiting upper bound can now be established.

Proposition 1. If  $x_i^*$  is the solution of the equation

$$\frac{(x^{1/2} + \eta_i)^2 \sigma_i^2}{(x^{1/2} + \eta_i)^2 + \sigma_i^2} = x,$$

then  $\lim_{n \rightarrow \infty} b_{ni}^2 = x_i^*$ .

Proof: Define

$$\phi(x) \triangleq \frac{(x^{1/2} + \eta_i)^2 \sigma_i^2}{(x^{1/2} + \eta_i)^2 + \sigma_i^2}$$

and note that  $\phi(b_{(n-1)i}^2) = b_{ni}^2$ . Result follows directly from application of Lemma 1. ||| Hence, use of the estimation algorithm in (23) yields estimates  $\{\hat{h}^k\}$  satisfying  $\lim_{k \rightarrow \infty} E \|\hat{h}^k - h^k\|^2 \leq \sum_{i=1}^p x_i^*$ .

#### LOWER BOUNDS ON THE ESTIMATION ERROR

We now derive a lower bound on the attainable mean-squared error, in the sense that with any linear estimator the error can always be as large as this bound if certain parameter changes occur. Of course, in many cases the actual error may be less than this lower bound, but we can never "guarantee" this fact. The basic idea is to construct a problem in which the system variations are random, with known mean and covariance, but bounded as in (10b). Kalman filtering techniques can then be applied in order to find the minimum mean-squared error which is attainable using any linear estimation procedure. This error provides the lower bound of interest.

Suppose that the  $w^k$ 's are random vectors, independent of the  $v^k$ 's, which satisfy (10b) and are described by

$$P[(w^k, \psi_i) = \eta_i] = \frac{1}{2} = P[(w^k, \psi_i) = -\eta_i], \quad (26)$$

and

$$E[(w^k, \psi_i)(w^l, \psi_j)] = \eta_i^2 \delta_{ij} \delta_{kl}.$$

Then

$$E w^k = 0, \quad E w^k (w^j)^T = Q \delta_{kj}, \quad (27)$$

where  $Q$  can be found from knowledge of the  $\{\psi_i\}$  and (26). Suppose also that  $h^1$  is a random vector with, say, mean zero and covariance matrix  $P_1 =$

$(X^T R^{-1} X)^{-1}$ . The error covariance matrix for the Kalman estimate  $\hat{h}^k$  of  $h^k$  is given by the recursive equation (see e.g., [4])

$$P_k = (P_{k-1} + Q)(I - X^T[X(P_{k-1} + Q)X^T + R]^{-1} \cdot X(P_{k-1} + Q)), \quad k = 2, 3, \dots \quad (29)$$

We then have:

Proposition 2. The upper bound on the attainable mean-squared error of any linear estimate of the system parameters  $\{h^n\}$ , in the model described by (9) and (10), cannot be less than  $\inf_k \text{Tr } P_k \triangleq e$ , where  $P_k$  ( $k \geq 2$ ) is given by (29) and  $P_1 = (X^T R^{-1} X)^{-1}$ .

Proof: Suppose the  $\{w^k\}$  are random vectors satisfying (26) and  $h^1$  is a random vector with mean zero and covariance  $P_1$ . First we observe that if only  $y^1$  is available, there is no linear estimator that will yield uniformly smaller mean-squared error than  $P_1$ . This is because  $P_1$  was chosen to be the variance of the LUMV estimate for  $h^1$  in the equation  $y^1 = Xh^1 + v^1$ , and no biased estimate gives a mean-squared error bounded for all  $h^1$ . Next, suppose that for  $m \geq 2$  there exists a linear estimator  $\hat{h}^m$  such that

$$E(\|\hat{h}^m - h^m\|^2 \mid w^1, \dots, w^{m-1}) < e$$

for every sequence  $\{w^1, \dots, w^{m-1}\}$ . Then

$$E\|\hat{h}^m - h^m\|^2 = E\{E(\|\hat{h}^m - h^m\|^2 \mid w^1, \dots, w^{m-1})\} < e.$$

But this contradicts the minimum mean-squared error property of the Kalman estimate  $\hat{h}^m$ . Consequently, for any  $m$ , there is a sequence  $w^1, \dots, w^{m-1}$  such that  $\inf_k \text{Tr } P_k \geq e$ . Since  $\{w^1, \dots, w^{m-1}\}$  satisfies (10b), this proves the assertion. |||

It is of interest to interpret this result in the special case of (9) and (10) where  $X$  is the identity and  $R$  is diagonal. The eigenvectors  $\{\psi_i\}$  then become the standard unit vectors in  $R^p$  and  $Q = \text{diag}(\eta_1^2, \eta_2^2, \dots, \eta_p^2)$ . Also, with a little algebraic manipulation (29) becomes

$$P_k = (P_{k-1} + Q)(P_{k-1} + Q + R)^{-1}R.$$

Since  $P_1 = R$  is diagonal, all the  $P_k$  will be diagonal. In particular, if  $P_{k-1} = \text{diag}(\alpha_{(k-1)1}^2, \alpha_{(k-1)2}^2, \dots, \alpha_{(k-1)p}^2)$  then  $P_k = \text{diag}(\alpha_{k1}^2, \alpha_{k2}^2, \dots, \alpha_{kp}^2)$  where

$$\alpha_{ki}^2 = \frac{(\alpha_{(k-1)i}^2 + \eta_i^2)\sigma_i^2}{\alpha_{(k-1)i}^2 + \eta_i^2 + \sigma_i^2}, \quad i = 1, \dots, p. \quad (30)$$

Corollary. If for the system described by (9) and (10),  $X$  is the identity and  $R$  is diagonal, then there exists a sequence  $\{w^k\}$  satisfying (10b) such that

$$E\|\hat{h}^k - h^k\|^2 \geq \frac{p}{2} \left\{ -\frac{\eta_1^2}{2} + \frac{\eta_1^2}{2} \sqrt{1 + \frac{\sigma_1^2}{\eta_1^2}} \right\}$$

for any linear estimate  $\hat{h}^k$  of  $h^k$ .

Proof: Define

$$\theta(x) \triangleq \frac{(x + \eta_i^2)\sigma_i^2}{x + \eta_i^2 + \sigma_i^2}$$

and note that for equation (30),  $\theta(\alpha_{(k-1)i}^2) = \alpha_{ki}^2$ . Application of Lemma 1 yields

$$\lim_{k \rightarrow \infty} \alpha_{ki}^2 = -\frac{\eta_i^2}{2} + \frac{\eta_i^2}{2} \sqrt{1 + 4\frac{\sigma_i^2}{\eta_i^2}}, \quad i = 1, \dots, p.$$

Since

$$\inf_k \text{Tr} P_k = \lim_{k \rightarrow \infty} \text{Tr} P_k = \lim_{k \rightarrow \infty} \sum_{i=1}^p \alpha_{ki}^2 = \sum_{i=1}^p \left\{ -\frac{\eta_i^2}{2} + \frac{\eta_i^2}{2} \sqrt{1 + 4\frac{\sigma_i^2}{\eta_i^2}} \right\},$$

the result follows directly from Proposition 2. ||| Figure 1 provides a plot of the  $i$ th component of the upper and lower bounds on the mean-square error when  $X = I$  and  $R$  is diagonal.

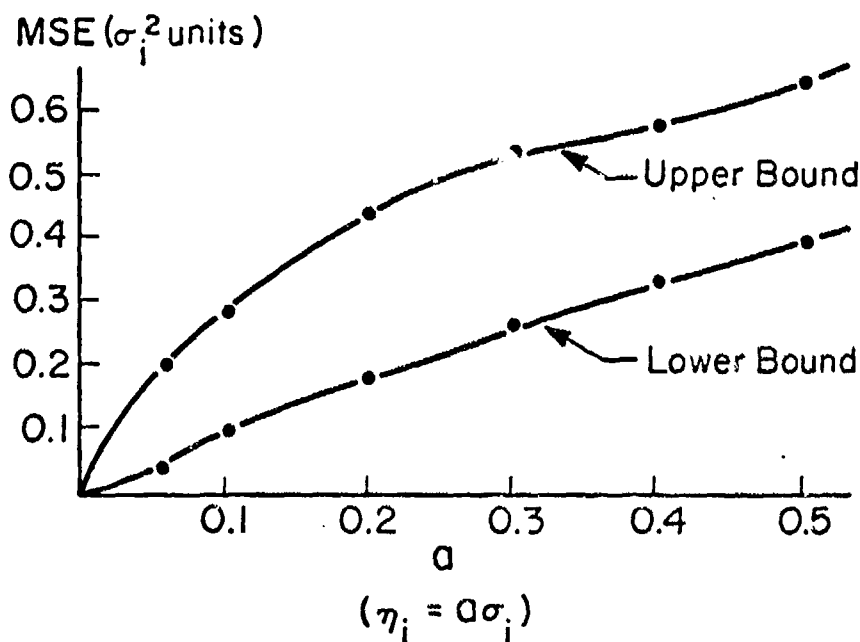


Figure 1. Bounds on Errors

#### COMMENTS

- (1) The results derived for estimating the parameters when they vary only between blocks of measurements can be extended to include the case where the parameters vary between each pair of measurements (see [3]).
- (2) The upper and lower bound expressions given by Propositions 1 and 2 provide criteria for determining whether or not a system can be considered as being slowly varying (see [3]).
- (3) The restriction to the case where  $X$  is the identity is in fact less restrictive than it would appear to be at first glance. The "standard

$\epsilon$ -representation" developed in [1] does provide such a model for a very large class of systems. Of course, the parametrization provided by this model may not be what is desired in a practical situation. However, our point of view as regards to the concept of slowly varying is that if there is any parametrization for which sufficiently good estimates can be made, then the system is slowly varying.

- (4) The estimate used at each state of the recursive estimation procedure is essentially the modified LUMV estimate described in [2].

#### REFERENCES

- [1] W. L. Root, "On the Modelling of Systems for Identification. Part 1.  $\epsilon$ -Representations of Classes of Systems. Part 2. Time-Varying Systems," to appear in SIAM J. Control.
- [2] W. L. Root, "On the Modelling and Estimation of Communication Channels," Multivariate Analysis III, ed. Krishnaiah, Academic Press, 1973.
- [3] P. H. Fiske and W. L. Root, "On the Modelling and Identification of Slowly Varying Systems," Technical Report, The University of Michigan, College of Engineering (in preparation).
- [4] K. J. Åström, Introduction to Stochastic Control Theory, Academic Press, 1970.